



Application of Principles of Testing: A Study in the Context of Third Year Engineering Graduates



Prativa Rani Saha ^a
R. Kalpana ^b

Article history:

Received: 25 November 2016

Revised: 10 December 2016

Approved: 22 December 2016

Published: 14 January 2017

Keywords:

Validity;

Accuracy;

Reliability;

Washback;

Language Testing;

Abstract

Language testing has been an elemental part of the teaching- learning the process, though views on language testing have always been changing. This paper presents the various components required to construct and use a language test. It primarily aims at highlighting the importance of validity, reliability, and washback and focuses if a test paper is valid, reliable and has a washback effect. A test should measure what it intends to measure and for this validity, of a test is an essential component. Reliability of a test refers to how a test should be measured and is also concerned with the consistency of scoring and the accuracy of the organization of the test, hence reliability also is an important factor for a test. Further, any item should ideally have a positive washback effect i.e., the test must have a beneficial effect. Any test setter has to consider these three qualities of a test discussed above. The paper aims at highlighting about these important aspects of test i.e. validity, reliability and washback and elucidates how these three aspects could be integrated into testing writing skills.

2455-8028 ©Copyright 2017. The Author.

This is an open-access article under the CC BY-SA license
(<https://creativecommons.org/licenses/by-sa/4.0/>)

All rights reserved.

Author correspondence:

Prativa Rani Saha,

PSG College of Technology, Coimbatore.

Research & Development Centre, Bharathiar University, Coimbatore

Email Address: prativa2007@gmail.com

1. Introduction

Many researchers discuss the issue of a paradigm shift in testing issues ranging from psychometrics (questionnaires, personality tests) to a broader model of assessment in education, from an examination culture to an assessment culture. Gipps (1990) also observes that a much wider range of assessment practices are now in use than there were twenty-five years ago. Teachers today in engineering usually design and administer classroom tests to their students for decision-making and problem-solving purposes. To aid in decision-making

^a VNR VJIET, Hyderabad

^b PSG College of Technology, Coimbatore

and problem-solving, it is imperative that these tests are both valid and reliable. Test validity and reliability could be achieved by taking a systematic approach to designing a test. The various components that are important for a test are validity, reliability, and washback and it is important that these principles are given their due consideration for a test to have a positive effect. This paper analyses a model question paper of a third year Engineering Graduate and examines if the paper is designed in a way to enhance the validity, reliability, and positive washback.

2. Research Methods

Definition and Principles of Assessment

Validity and Reliability are two principle components of a test that indicate the quality and usefulness of the test. Hence, while evaluating the suitability of the test we have to examine these two features. Validity is the extent to which a test measures what it is supposed to measure. It also refers to what characteristic the test measures and how well the test measures that particular characteristic. This principle also demonstrates how good a test is for a particular situation. The question of validity is raised in the contexts where we have to decide the form of the test, the purpose of the test and the population for whom the test is intended. Thus, in place of the routine question “Is this a valid test?”, the question “how valid is this test for the decision that I need to make?” needs to be considered. Validity testifies if the characteristic being measured by a test is related to job qualifications and requirements or if there exists any link between the test performance and job performance. For example, if a test is based on to be a valid predictor of the performance for a specific job, and then a person who scores high in the test is also likely to perform well in the job. Validity also describes the degree to which specific conclusions about the test-takers can be made. In other words, it indicates the usefulness of the test.

According to [Nunally \(1982\)](#), reliability is concerned with the extent to which measurements are repeatable if all items being studied were included. In other words, reliability can be described as the extent to which a test measures what it aims to measure consistently and accurately. Reliability is the degree to which a test consistently measures whatever it measures. Errors of measurement that affect reliability are random errors and the errors of measurement that affect validity are systematic or constant errors. (Gay) A test is considered reliable if it yields the same results each time the test is conducted. For example, if we take a test in synonyms today to assess whether we remember the meanings of the lesson and are able to use them and then took another test on the same next week, we would expect to see similar scores on both the tests. This would mean the assessment was reliable. Reliability in an assessment is important because assessments provide information about student achievement and progress. The reliability principle is, therefore, concerned with whether a test gives consistent results, as Underhill (1987, p. 9) says: If the same learners are tested on two or three occasions, do they get the same score each time? Furthermore, the terms test-retest and parallel tests were emphasized by teachers when discussing the reliability concept even though those methods have problems. A test must be reliable, as a test cannot be valid unless it is reliable.

The third principle of assessment is washback which also has a lot of importance while designing a paper as the paper must have a beneficial effect. Washback has led to a greater understanding of this construct than was available earlier. Washback in simple terms refers to the impact the test has on teaching and learning. Moreover, when the concept of washback was examined, it was clear that it is a powerful tool for both language learners and teachers. The final principle, washback, “refers to the effects the tests have on instruction in terms of how students prepare for the test” ([Brown & Abeywickrama, 2010, p. 38](#)). The literature suggests that the focus should be on impact and not processes. The washback effect of testing is a well-designed phenomenon common to all. The washback or backwash mainly emphasizes on the effect that the test would have. As rightly said: ‘the influence of testing on teaching and learning’ ([Gates, 1995](#)). We as teachers, always design our tests to keep it outcome based. As G. Buck rightly observes, ‘There is a natural tendency for both teachers and students to tailor their classroom activities to the demands of the test, especially when the test is very important to the future of the students, and pass rates are used as a measure of teacher success. This influence of the test on the classroom (referred to as washback by language testers) is, of course, very important; this washback effect can be either beneficial or harmful.’ Hence, the washback effect of a test is decided by the test focus, test type, test content and test delivery. The test focus describes whether the test has a larger or narrower scope which should be decided while designing a test paper. Test type also focuses on whether the test is Computer based, Paper-based, Oral Proficiency test etc. While the test content states whether the test has a positive effect of the syllabus. Test delivery explains whether the test is from textual to oral or other modes. There could be a positive or negative washback of the test. Hence, studying the washback effect of tests is also a very valuable as it helps in

verifying the appropriateness of the test. So, tests must be part of learning experiences for all involved. In a later paper, Shohamy (1993a, p. 4) summarized four key definitions that are useful in understanding the washback concept:

- a) Washback effect refers to the impact that tests have on teaching and learning.
- b) Measurement-driven instruction refers to the notion that tests should drive learning.
- c) Curriculum alignment focuses on the connection between testing and the teaching syllabus.
- d) Systemic validity implies the integration of tests into the educational system and the need to demonstrate that the introduction of a new test can improve learning.

In this paper, a 3rd Year, External Lab examination paper of an Engineering College is analyzed and the common characteristics of the test samples are scrutinized in terms of validity, reliability, and washback effect. The syllabus of a third year Engineering consists of Speaking, Reading and Writing Skills and all the three skills were tested. Writing skills include topics such as writing a SOP, abstract, report, summarizing & synthesizing, and writing a resume and cover letter. The question paper requires students to write questions based on writing for an academic and specific purpose. The sample question paper of a 3rd Year paper is analyzed against the three principles is as follows:

III B. TECH –I SEM. ADVANCED ENGLISH COMMUNICATION SKILLS LAB EXTERNAL EXAMINATION

Time: 3 hrs

Max. Marks: 70

Section- I (Speaking Skills)

35 Marks

I. Participate in the Group Discussion

20 Marks

II. Participate in JAM.

15 Marks

Section- II

Part-A (Reading and Writing Skills)

15 Marks

1. Read the abstract given below and use the information to write the **Introduction, Discussion & Recommendations** for a report.

Renewable Energy Resources

The electricity requirements of the world including India are increasing at an alarming rate and the power demand has been running ahead of supply. It is also now widely recognized that the fossil fuels and other conventional resources, presently being used for generation of electrical energy, may not be either sufficient or suitable to keep pace with ever increasing demand of the electrical energy of the world. The recent severe energy crisis has forced the world to develop new and alternative methods of power generation, which could not be adopted so far due to various reasons. The magneto-hydro-dynamic (MHD) power generation is one of the examples of a new unique method of power generation. The other non-conventional methods of power generation may be such as solar cells, fuel cells, thermoelectric generator, thermionic converter, solar power generation, wind power generation, geothermal energy generation, tidal power generation etc. This paper elucidates about Different Energy sources, why we are going for non-conventional energy sources, Different non-conventional energy sources & comparison between them, about fuel cells and their applications.

2. You are applying to a University to pursue Master Degree (MS) in your specific field of study. Write a **Statement of Purpose** as part of the application process.

Section- III (Behavioural and Interview Skills)

Marks: 2X10=20

3. Respond to the advertisement given below and draft your resume and cover letter for the same.
 - 1] Wanted Junior Engineers (Projects), with B.Tech./B.E, and must have 1-2 years of experience. The work responsibility includes:
 - [a] Operate software like Autocad, Ms. Office.
 - [b] Create a schedule and prepare evaluation/monitoring reports.
 - [c] Supervise the development of work.
 - [d] Able to work independently.

4. You are the student chapter chairman and you want a student- activity- center to be built in the college premises for ECA/CCA activities. You have a meeting with your Principal to discuss the same. Use your negotiation and persuasion skills to convince him about the requirement. Draft the dialogue.
5. Write an **abstract** on a topic that you have presented in your college technical fest.

The question paper includes writing a sop, report, abstract, dialogue, resume and letter. A test is valid when it assesses the objectives and what has been taught. The paper should be designed so that it meets the objectives. The course outcomes easily align objectives with course content and evaluation methods and establish a logical sequence of learning milestones. Hence the objectives and outcomes of the Course are:

Course Objectives:

- 1] Enable the students to create clear, accurate, and succinct content to write business letters, resume, SOP, Proposals and Technical Reports for academics as well as for workplace
- 2] To adjust technical content to meet the needs of a specific target audience
- 3] Groom students to speak accurately and fluently and prepare them for real-world activities through behavioral skills.
- 4] Train students in soft skills through role play and group discussion to improve their EQ.

Course Outcomes:

Students will be able to:

- 1] Comprehend various genres of technical writing generated/encountered in the engineering profession.
- 2] Summarize and synthesize information and produce technical writing that is required in academics as well as in the engineering profession.
- 3] Write covering letters, resume, SOP, Project Proposals and Technical Reports
- 4] Speak fluently and address a large group of audience and participate in debates and discussions.
- 5] Negotiate terms, manage complex situations through interpersonal skills, persuade people and make quick decisions.

The elements tested here are content, format, style and organization. The questions included in the test intend to test what it proposes to measure, i.e. if the students were able to write a sop or an abstract as they were taught and in case if they have followed the steps. It also offers meaningful information about the students' ability whether they are capable of writing the given assignment.

To ensure reliability, the instructors have provided clear administration and scoring instructions. The questions listed above can be considered to be reliable because the students will be able to write a sop, cover letter, abstract during an exam and also in different situations while applying for a job interview or while applying for admission into higher studies. The rubrics for reliability are uniform. It is also important that the other test factors such as the formats and content of the questions and the length of the exam are consistent. Hence, the test is reliable as the students will be able to write the given topics in different circumstances too.

It refers to the outcomes for the learner, the teacher, and the teaching context (Linville). To examine if the test has a washback effect, it should be:

- 1] Be more formative in nature than summative.
- 2] Give chance to learners to prepare.
- 3] Positively impact what teachers teach, and what learners learn etc.

Washback motivates the students to learn more. The topics as sop, abstract and report will surely create an interest in students to learn and know more about them as they must be able to produce an impressive sop for the university or an abstract for a presentation. The students realize the importance of these and the test definitely has a positive washback.

The test item also examines if learners are able to analyze, infer, interpret and reason it out to attempt the answer. Learners taking such tests are also required to organize ideas, use linkers and sequence ideas logically to construct paragraphs. The question paper provides a challenge to the learners to apply what they have learnt in the classroom. Thus, the questions contain ample scope for learners to analyze, organize and construct coherent passages for letter, abstract, sop and report. Hence, the question paper was designed in context with the principles of validity, reliability, and washback in long run.

3. Results and Analysis

The course outcome attainment based on these questions and the result obtained after the test show that the question paper is designed based on the principles of assessment. It has been observed that there are some issues to be taken into consideration when designing language exams. The exam papers collected for the purpose of this study reveal the following:

- a) Around 60 students appeared for the test and about 75% were able to perform well in the exam.
- b) The paper was designed to match with the principles of assessment in mind. As a good number of students performed in the test, it also confirms that the test was valid, reliable and had washback as it has a positive impact, a good number of students performed well and was able to take it to the next level etc.
- c) Having examined these items the result proves that the paper was designed well.

4. Conclusion

The past ten years have seen a growing awareness that testing can have consequences beyond just the classroom. Tests and test results have a significant impact on the career or life chances of individual test takers (e.g. access to educational/employment opportunities). The presented analysis shows that the test discussed here had reliability, validity, and washback issues as it has uniform rubrics for scoring, measures what it exactly proposes to measure and offers a chance to learners to prepare and gives feedback relevant to the objective. As stated by Morrow, the test reflects and encourages good classroom practice. If employed this way testing becomes a method parallel to teaching and learning and the teaching-learning environment becomes more conducive.



Acknowledgments

My deep and sincere gratitude to my Supervisor, Dr. R. Kalpana, Asst. Prof., PSG College of Technology for her support, suggestion and contribution in finishing this research paper. I am thankful to my friends for their suggestions to complete the paper. Last but not the least, I also thank my family for supporting me to complete my work. I also thank IJCU of International Journal for publishing my paper.

References

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford University Press.
- Brown, J. D. (2002). Extraneous variables and the washback effect. URL: http://jalt.org/test/bro_14.htm (дата обращения: 13.12. 2013).
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment* (Vol. 17). Sage publications.
- Davidson, F., & Lynch, B. K. (2008). *Testcraft: A teachers guide to writing and using language test specifications*. Yale University Press.
- Gay, L. R. (1987). *Professional Supplement for Educational Research: Competencies for Analysis and Application*. Merrill Publishing Company.
- Guilford, J. P. (1954). Psychometric methods.
- Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language testing*, 14(3), 295-303.
- Harmer, J. (2001). *The practice of English language teaching*. London/New York.
- Heaton, J. B. (1990). *Writing English Language Tests New Edition* Consultants editor Jeremy Harmer and Roy Kingsbury.
- Weir, C. J.(2005). *Language testing and Validation*. UK: Palgrave Macmillan.
- Winer, B. J. (1991). *Statistical principles in experimental design* (No. 519.5 W55).

Biography of Authors

	<p>Ms. Prativa Rani Saha has been working as Assistant Professor in the Department of Humanities & Sciences, VNR Vignana Jyothi Institute of Engineering & Technology since 2008. She is currently pursuing her Ph. D from Bharathiar University, Coimbatore. Her research interests include ELT, Academic Writing, Literature, Task Based Teaching Techniques and Educational Technology. She has completed her PGDTE from EFLU, Hyderabad.</p>
	<p>Dr. R. Kalpana works as an Assistant Professor in the Department of English, PSG College of Technology since 1998. Her research interests include ELT, Linguistics, Academic Writing Stylistics and Communication. She took her Ph.D. from the University of Madras and her PGCTE from EFLU, Hyderabad.</p>